Joint Pairing and Structured Mapping of Convolutional Brain Morphological Multiplexes for Early Dementia Diagnosis

Alzheimer's Disease Neuroimaging Initiative; Lisowska, Anna; Rekik, Islem

Published in:
Brain Connectivity

DOI:
10.1089/brain.2018.0578

Publication date:
2019

Document Version
Peer reviewed version

Joint Pairing and Structured Mapping of Convolutional Brain Morphological Multiplexes for Early Dementia Diagnosis

Anna Lisowska, Islem Rekik*†, and for the Alzheimer's Disease Neuroimaging Initiative‡

BASIRA lab, CVIP group, School of Science and Engineering, Computing, University of Dundee, UK

Correspondence and requests for materials should be addressed to I. Rekik (irekik@dundee.ac.uk)

Address: Queen Mother Building

University of Dundee

BASIRA Lab (www.basira-lab.com)

DD1 4HN Scotland, UK

Office: Office 2.18 QMB

Tel: +44 (0)1382 384908

Email: irekik@dundee.ac.uk

Keywords: morphological brain network, cortex morphology, convolutional brain multiplex, canonical correlation analysis, ensemble classifier, early dementia diagnosis

† Corresponding author.

‡ Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Abstract

Diagnosis of brain dementia, particularly early mild cognitive impairment (eMCI), is critical for early intervention to prevent the onset of Alzheimer's Disease (AD), where cognitive decline is severe and irreversible. There is a large body of machine-learning based research investigating how dementia alters brain connectivity, mainly using structural (derived from diffusion MRI) and functional (derived from resting-state functional MRI) brain connectomic data. However, how early dementia affects cortical brain connections in *morphology* remains largely unexplored. To fill this gap, we propose a *joint morphological brain multiplexes pairing and mapping strategy* for early MCI detection, where a brain multiplex not only encodes the relationship in *morphology* between pairs of brain regions, but also a pair of brain morphological networks. Experimental results confirm that the proposed framework outperforms in classification accuracy several state-of-the-art methods. More importantly, we unprecedentedly identified most discriminative brain morphological networks between eMCI and NC, which included the paired views derived from maximum principal curvature and the sulcal depth for the left hemisphere and sulcal depth and the average curvature for the right hemisphere. We also identified the most *highly correlated morphological brain connections* in our cohort, which included the (pericalcarine cortex, insula cortex) on the maximum principal curvature view, (entorhinal cortex, insula cortex) on the mean sulcal depth view, and (entorhinal cortex, pericalcarine cortex) on the mean average curvature view, for both hemispheres. These highly correlated morphological connections might serve as biomarkers for early MCI diagnosis.

1. Introduction

Early diagnosis of brain dementia, specifically mild cognitive impairment (MCI), which might progress towards Alzheimer's disease (AD), might help prevent the onset of AD through early efficient intervention. There is evidence that MCI alters brain morphology, including cortical thinning (Querbes *et al.*, 2009). For instance, previous research (Im *et al.*, 2008) found that sulci in brains of MCI patients were characterized by reduced curvature, with sulcal widening observed and sulcal depth reduced compared to controls. (Liu *et al.*, 2012) demonstrated decreased global sulcal index and increased widths of nearly all individual sulci in MCI, while (Hamelin *et al.*, 2014) found that cortical thickness, the hippocampal volume and the sulcal width to be the best markers for distinguishing MCI from NC. This is of great clinical value as it might help individualize early intervention to effectively alleviate the symptoms of the disease (Prince *et al.*, 2013).

This early stage of AD was shown to affect functional and structural brain connectivities (obtained from functional MRI (fMRI) and diffusion-weighted MRI (dMRI)) (Bullmore and Sporns, 2009), thereby causing cognitive decline not fulfilling the AD criteria, but greater than expected of their age and educational level (Mucke, 2009). Recently, more research focused on accurate detection of early MCI (eMCI), which is essential for slowing down potential conversion to AD. For instance, (Prasad *et al.*, 2015) investigated the predictive power of various combinations of connectomic features, such as pairwise connectivity and maximum flow between two brain regions, extracted from dMRI images for eMCI and normal control (NC) classification problem. More recently, (Wee *et al.*, 2016) computed sparse temporal networks using sliding-

window approach over a time series of resting-state functional MRI. (Chen *et al.*, 2016) extended this work by additionally considering the high-order correlation between different pairs of brain regions. By combining low-order and high-order brain networks, they further improved the classification accuracy of eMCI/NC patients.

Other studies showed that properties of networks, constructed by volumetric (Yao *et al.*, 2010) and geometric morphological measures (Li *et al.*, 2016), were affected by MCI. However, all these connectomic studies relied on using functional or diffusion-based MRI, which are much more difficult to acquire as they are time-consuming, costly and prone to noise, and are not conventionally used in the diagnostic routine. According to the tension theory of cerebral cortex morphogenesis, network changes in the morphological attributes (surface) of the brain reflect the underlying changes in the structural and functional connectivity (Van Essen, 1997) and can be studied without the need for costly and time-consuming imaging of patients. Hence, we recently introduced the use of *morphological* brain network for eMCI diagnosis (Lisowska and Rekik, 2017). Specifically, we devised an ensemble classifier architecture leveraging a novel representation of multi-layer morphological cortical networks for dementia onset identification.

Previous research on dementia state classification showed that using multi-layer networks (i.e., stacking different networks) improved the prediction accuracy for disease identification when compared to using single view networks (Giuliano Zippo and Castiglioni, 2016; La Rocca *et al.*, 2017; Crofts *et al.*, 2016). However, none of these multi-layer network-based methods explored the relationship between two consecutive layers in the network or cortical morphology (Brown and Hamarneh, 2016).

To fill this gap, we proposed a multi-layer network (multiplex), consisting of multiple morphological brain network views (Lisowska and Rekik, 2017). We note that a simple concatenation of multiple networks hinders the investigation of potentially complex changes in cortical regions, which might vary jointly or independently across different brain views as they become affected by dementia onset. Hence, we introduce inter-layers into a multiplex structure to capture the relationship between different brain views. Since each multiplex is not invariant to the ordering of the intra-layers, in our previous research (Lisowska and Rekik, 2017), we generated multiple multiplexes for each subject while considering all possible combinations of intra-layers, thereby capturing all relationships between different brain views. However, this resulted in a highly correlated data and many redundant features. To address this limitation, in this work we propose a new shallow multiplex structure, each consisting of two morphological views with a single inter-layer between them.

Next, to leverage complementary information from different brain multiplexes, we previously used canonical correlation analysis (CCA) to map two sets of multiplex features into a shared space where they become more comparable (Lisowska and Rekik, 2017; Haghighat *et al.*, 2016; Zhu *et al.*, 2016). However, one of the main limitations of CCA is that it lacks biological interpretability, as it does not perform feature selection (Chen *et al.*, 2013). Sparse CCA was shown to solve this issue by computing the relationship between modalities using much less features (Parkhomenko *et al.*, 2009). Since connections in the brain tend to be affected jointly by a disease, we propose to embed the recently developed Structured Sparse CCA (SS-CCA) (Du *et al.*, 2017) for morphological multiplex fusion into our eMCI/NC classification framework.

Leveraging the strengths of ensemble classifier learning (Džeroski and Ženko, 2004; Quan *et al.*, 2016)., we propose structured ensemble classifier learning using

multiple sets of shallow brain multiplexes, where each pair of multiplex sets is mapped onto a SS-CCA space then fused. Ultimately, we use the fused multiplex features to train a linear classifier in each spanned SS-CCA space for early MCI identification.

2. Methods and Materials

2.1 Proposed Joint Pairing and Structured Mapping Strategy using

Fig. 3 displays the key steps of the proposed framework, where shallow brain multiplexes illustrated in **Fig. 2**, are used to train an ensemble classifier architecture via a data pairing strategy. First, each subject is represented by a set of brain multiplexes. Each colored cube denotes a unique multiplex constructed from two brain networks (e.g., derived from cortical thickness and sulcal depth). Second, we use PCA to reduce the dimensionality of all training multiplexes. Third, we map each pair of training multiplexes onto a space where their correlation is maximized. Next, each mapped pair from two multiplex sets will serve as a training sample to learn an SVM classifier. Last, using a weighted majority voting by all classifiers we identify the label of a new testing subject.

Convolutional Morphological Brain Multiplexes

In this section, we introduce the concept of a d-layer ($d \geq 2$) convolutional brain multiplex and present our novel *structured* mapping strategy using paired sets of brain multiplexes. Matrices are denoted by boldface capital letters, e.g., $\boldsymbol{X}$, and scalars are denoted by lowercase letters, e.g., x. We denote the transpose operator and the trace operator as $\boldsymbol{X^T}$ and $tr(\boldsymbol{X})$, respectively. For easy reference and enhancing the readability, we have summarized the major mathematical notations in **Table 1**. We

illustrate in **Fig. 2** the proposed framework for shallow (2-layer) brain multiplex construction and in **Fig. 3** pairing-based ensemble classifier learning using SS-CCA mapping of sets of brain multiplexes.

Table 1: *Major mathematical notations used in this paper.*

Mathematical notations	Definition
$\boldsymbol{V}$	brain network (single view) in $\mathbb{R}^{n\times n}$
$\boldsymbol{M}$	brain multiplex composed of intra-layers and convolutional inter-layers
$\boldsymbol{C}_{i,j}$	convolutional intra-layer between consecutive brain network views $\boldsymbol{V}_i$ and $\boldsymbol{V}_j$ in $\boldsymbol{M}$
$\mathbb{M} = \{\boldsymbol{M_1}, \ldots, \boldsymbol{M_N}\}$	subject-specific brain multiplexes with diff erent orderings of intra-layers
$\boldsymbol{M_k}$	matrix in $\mathbb{R}^{d\times N_s}$ containing the d multiplex features for all N_s training samples from multiplex $\boldsymbol{M_k} \in \mathbb{M}$
$\boldsymbol{M_{k,l}} = [\boldsymbol{M_k}, \boldsymbol{M_l}]$	paired multiplex feature matrices derived from two training multiplexes in $\mathbb{M}$
$\boldsymbol{G}_{\mathbf{M}_k,\mathbf{M}_l}$	graph modeling the relationship between features extracted from multiplexes $\boldsymbol{M_l}$ and $\boldsymbol{M_k}$
u	learned sparse SS-CCA weighting vector for $\boldsymbol{M_k}$-derived features
v	learned sparse SS-CCA weighting vector for $\boldsymbol{M_l}$-derived features

Single-view morphological network construction. In line with the works of (Mahjoub et al., 2018) and (Soussia and Rekik, 2017), we define morphological brain networks as follows. For each cortical attribute (e.g., cortical thickness), we construct a single-view network for each subject. Such network comprises a set of nodes (anatomical brain regions) and a collection of edges interconnecting the nodes (representing the disimilarity between the two brain regions in morphology). The average value of a cortical attribute was calculated for each anatomical region of interest (ROI). For each cortical attribute, the strength of each network edge connecting two ROIs is then computed as the absolute difference between their average values, thereby quantifying their dissimilarity **(Fig. 1**). The same procedure was followed to

obtain the connectivity matrices from different cortical attributes (e.g., sulcal depth, curvature). We note that a morphological brain connectivity models disimilarity in morphology between anatomical brain regions (similarly to functional connectivity, which models correlation between firing neurons), rather than being a real physical connection (like structural connectivity). We believe that both functional and morphological connections mediate 'real' connections., as there is a relationship between brain function, morphology and structure (Van Essen, 1997).

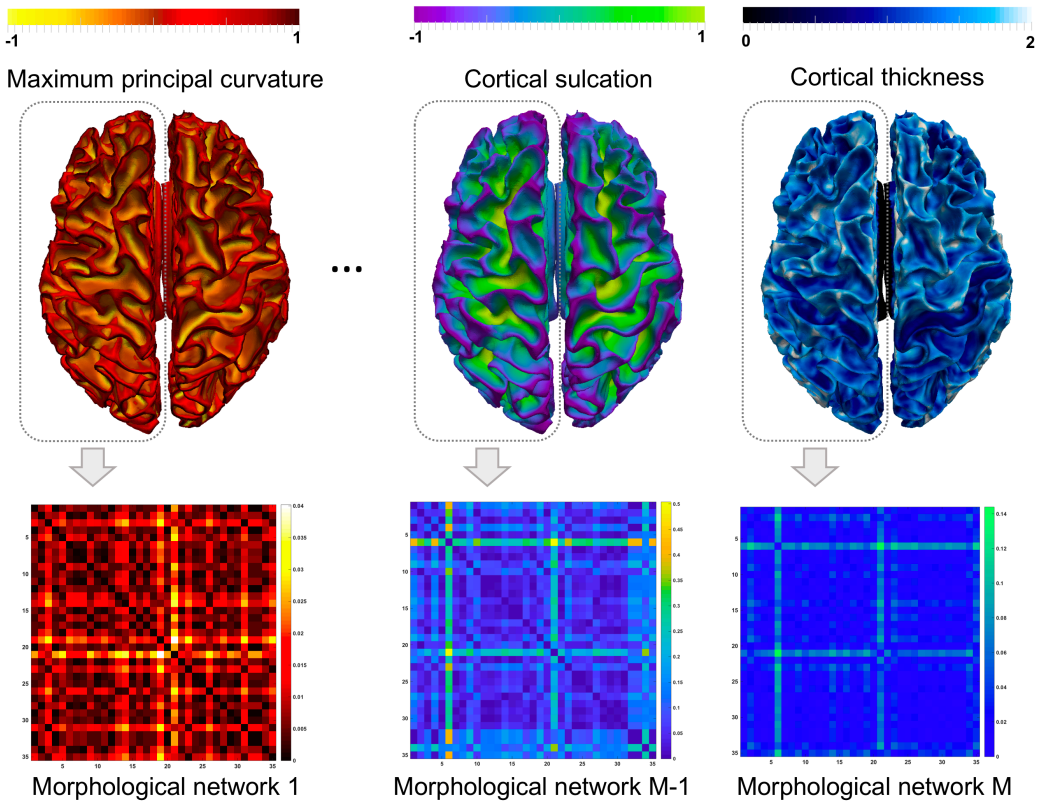


Figure 1: *Morphological brain network estimation using different cortical attributes.*

Convolutional brain multiplex construction. In a generic way, we define a brain multiplex $\boldsymbol{M}$ using a set of M intra-layers $\{\boldsymbol{V_1}, \ldots, \boldsymbol{V_M}\}$, each representing a single view of the brain morphology (i.e., cortical attribute), where between two consecutive intra-layers $\boldsymbol{V_i}$ and $\boldsymbol{V_j}$ we slide an inter-layer $\boldsymbol{C_{i,j}}$, which is defined by convolving two consecutive intra-layers. The convolutional inter-layer models the relationship between

two layers. More specifically, the convolution 'blends' two layers together and is the expression of the amount of overlap of one layer as it is shifted over another. Each element in row a and column b within the convolutional inter-layer matrix $\boldsymbol{C_{i,j}}$ between views $\boldsymbol{V_i}$ and $\boldsymbol{V_j}$ is defined as: $\mathbf{C}_{i,j}(a,b) = \sum_p \sum_q \boldsymbol{V_i}(p,q)\boldsymbol{V_j}(a-p+1, b-q+1)$. The multiplex architecture allows not only to explore how different brain views get altered by a specific disorder, but how their relationship might get affected. Since the morphological brain connectivity matrices are symmetric (**Fig. 1**), we extract features from each multiplex by directly concatenating the off-diagonal weights of all connections in each triangular matrix. For each network of size $n \times n$, we extract a feature vector of size $(n \times (n-1)/2)$. Previously, in (Lisowska and Rekik, 2017), we introduced the generalized multiplex architecture: $\boldsymbol{M} = \{\boldsymbol{V_1}, \boldsymbol{C_{1,2}}, \boldsymbol{V_2}, \ldots, \boldsymbol{V_i}, \boldsymbol{C_{i,j}}, \boldsymbol{V_j}, \ldots, \boldsymbol{V_M}\}$. We note that, for a specific multiplex, we were only allowed to explore similarities between consecutive layers. To explore the inter-relationship between *all possible* combinations of intra-layers, we generated for each subject *N* multiplexes through simply reordering the intra-layer networks, thereby generating an ensemble multiplexes $\mathbb{M} = \{\boldsymbol{M_1}, \ldots, \boldsymbol{M_N}\}$ (**Fig. 2**). However, this approach resulted in many highly correlated multiplex features used for the ensemble classifier learning, which may somewhat mislead CCA mapping. To minimize the correlation between different multiplexes, we propose a shallow (i.e., 2-layer) convolutional multiplex structure. We define a shallow multiplex $\boldsymbol{M} = \{\boldsymbol{V_i}, \boldsymbol{C_{i,j}}, \boldsymbol{V_j}\}$ using 2 intra-layers $\boldsymbol{V_i}$ and $\boldsymbol{V_j}$ and an inter-layer $\boldsymbol{C_{i,j}}$ encoding the relationship between $\boldsymbol{V_i}$ and $\boldsymbol{V_j}$, slid in between them (**Fig. 2**). We note that each subject-specific brain multiplex in $\mathbb{M}$ captures unique similarities between 2 different morphological brain network views (e.g., sulcal depth network and cortical thickness network) that are not present in a different shallow multiplex.

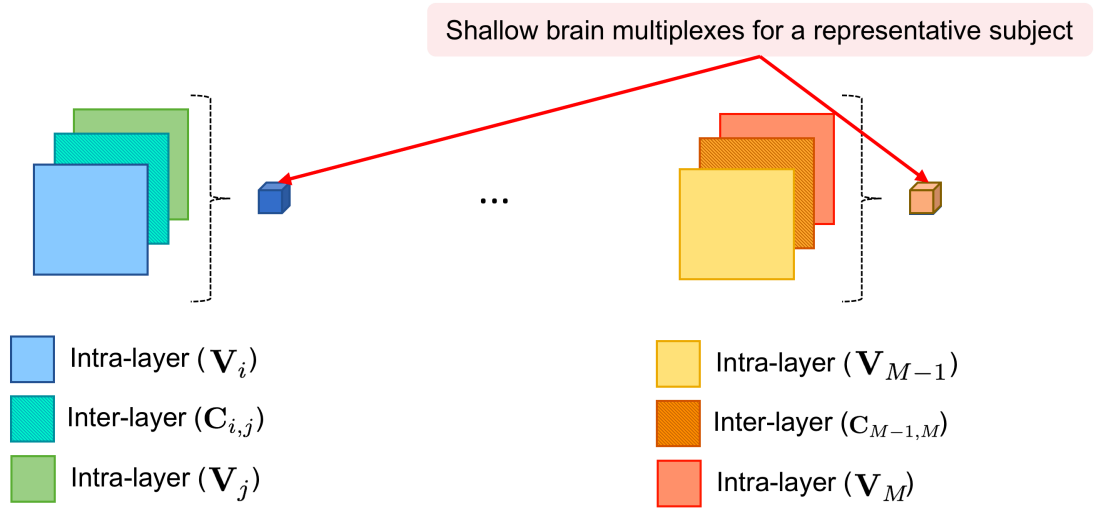


Figure 2: *Proposed shallow (2-layer) morphological brain multiplex construction for a single subject.* For M views (or networks), we generate C_M^2 brain multiplexes by enumerating the number of pairs of views.

Proposed canonical correlational mapping of brain multiplex sets. Since each multiplex $\boldsymbol{M_k} \in \mathbb{M}$ captures a unique and complex relationship between different brain network views, one needs to examine all morphological brain multiplexes in the ensemble $\mathbb{M}$. This will provide us with a more holistic understanding of how explicit morphological brain connections can be altered by dementia onset as well as how their implicit high-order (a connection of connections) relationship can be affected. To make use of all the information available from different multiplexes, we seek a feature fusion method that would extract the most relevant features for the classification task, minimize the modality-specific noise and reduce data dimensionality. Canonical Correlation analysis (CCA) was shown to be efficient in analysing and fusing associations between two sets of variables (Zhu *et al.*, 2016; Haghighat *et al.*, 2016) by identifying the structure common to the 2 views and creating a subspace that is robust to noise from different modalities. Since, CCA aims to find such subspace where correlations between projected features are maximized, and the noise present in either modality that is uncorrelated with the other modality is suppressed in the projected

subspace (Singanamallia *et al.*, 2014). However, due to the high dimensionality of our multiplex data, the CCA cannot be used without previously applying a dimensionality reduction technique to the data due to the high-dimensional covariance matrix singularity (Chen *et al.*, 2013). Therefore, we first apply Principal Component Analysis (PCA) to our data prior to CCA-based fusion step. PCA was shown to positively affect the classification performance when using morphological features for MCI diagnosis (Park *et al.*, 2012).

Proposed Structured Sparse CCA (SS-CCA) mapping of pairs of shallow brain multiplexes. Compared to conventional CCA mapping, the Structured Sparse CCA mapping brings two main advantages. First, it performs feature selection by imposing a sparsity constraint on the linear coefficients (Parkhomenko *et al.*, 2009), which means the correlation between the modalities is computed using much less features, and without the need to previously apply a dimensionality reduction. Second, it imposes sparsity in a structure-aware manner to capture the high-level structure information (Chen *et al.*, 2013), which is of great importance when extracting features from brain network data. In this paper, we particularly use graph guided pairwise group lasso (GGL) based sparse canonical correlation analysis model (GGL-SCCA), recently introduced by (Du *et al.*, 2017), as it has many advantages compared to methods using the group lasso or the graph/network guided fused lasso penalty to find the group structure. First, GGL-SCCA can recover the structure information from the input data in an unsupervised manner and without any a priori knowledge. Second, it has a strong upper bound for the grouping effect of correlated variables independently of sample correlation sign. Third, the GGL-SCCA finds stronger, more stable canonical correlations and cleaner canonical loading patterns comparing to other state-of-the-art SS-CCA methods (Chen and Liu, 2012; Du *et al.*, 2016). Specifically, we model the

relationship between features extracted from two multiplexes $\boldsymbol{M_k}$ and $\boldsymbol{M_l}$ using a graph $\boldsymbol{G_{\mathbf{M}_k,\mathbf{M}_l}} = (\boldsymbol{E_{\mathbf{M}_k,\mathbf{M}_l}}, \boldsymbol{V_{\mathbf{M}_k,\mathbf{M}_l}})$. We then identify the most correlated features through estimating the weighting vector $\mathbf{u}$, which sparsifies the edge set in $\boldsymbol{E_{\mathbf{M}_k,\mathbf{M}_l}}$ in $\boldsymbol{G_{\mathbf{M}_k,\mathbf{M}_l}}$ through solving the following Graph Guided Pairwise Group Lasso (GGL):

$$\Omega_{\mathrm{GGL}}(\mathbf{u}) = \sum_{(i,j)\epsilon \boldsymbol{E_{\mathbf{M}_k,\mathbf{M}_l}}} \sqrt{u_i^2 + u_j^2} \qquad (1)$$

GGL computes the Euclidian distance between features and encourages correlated features to be assigned similar canonical weights. This way GGL-SCCA overcomes the limitations of Sparse CCA, which follows an assumption that all the features within the same view are independent of one another. Because GGL uses sample correlation to define the group constraint, unlike previous SS-CCA methods, it does not require prior knowledge about the group structure, which may be unavailable or incomplete for the biomedical data. Ultimately, we formulate GGL-SCCA using pairs of brain multiplexes as follows:

$$\begin{cases} \min_{\mathbf{u},\mathbf{v}} -\mathbf{u}^{\mathrm{T}}\mathbf{M_k}^{\mathrm{T}}\mathbf{M_l}\mathbf{v} \\ s.t. \left|\left|\mathbf{M_k}\mathbf{u}\right|\right|^2 \leq 1, \left|\left|\mathbf{M_l}\mathbf{v}\right|\right|^2 \leq 1, \Omega_{\mathrm{GGL}}(\mathbf{u}) \leq c_1, \Omega_{\mathrm{GGL}}(\mathbf{v}) \leq c_2 \end{cases} \qquad (2)$$

And can be solved using the Lagrange method as in (Du *et al.*, 2017):

$$\mathcal{L}(\mathbf{u},\mathbf{v}) = -\mathbf{u}^{\mathrm{T}}\mathbf{M_k}^{\mathrm{T}}\mathbf{M_l}\mathbf{v} + \frac{\alpha_1}{2}\left|\left|\mathbf{M_k}\mathbf{u}\right|\right|^2 + \frac{\alpha_2}{2}\left|\left|\mathbf{M_l}\mathbf{v}\right|\right|^2 + \lambda_1\, \Omega_{\mathrm{GGL}}(\mathbf{u}) + \lambda_2 \Omega_{\mathrm{GGL}}(\mathbf{v})$$

Where the closed form solution is for each iteration t:

$$\begin{cases} \mathbf{u}^{t+1} = (\lambda_1 \mathbf{D}_1^t + \alpha_1 \mathbf{M_k}^T \mathbf{M_k})^{-1} \mathbf{M_k}^T \mathbf{M_l} \mathbf{v}^t, \\ \mathbf{v}^{t+1} = (\lambda_2 \mathbf{D}_2^t + \alpha_2 \mathbf{M_l}^T \mathbf{M_l})^{-1} \mathbf{M_l}^T \mathbf{M_k} \mathbf{u}^t \end{cases} \qquad (3)$$

Where $\mathbf{D}_1$ denotes the diagonal matrix with the k_1-th element being $\sum_{i,i\neq k_1} \frac{1}{\sqrt{u_{k_1}^2+u_i^2}}$ $(k_1 \in [1,p])$ and $\mathbf{D}_2$ denotes the diagonal matrix with the k_2-th element being $\sum_{j,j\neq k_2} \frac{1}{\sqrt{v_{k_2}^2+v_j^2}}$ $(k_2 \in [1,q])$, α and λ are tuning parameters. Both p and q values denote the upper bound for parameters k_1 and k_2, respectively.

The grouping effect of GGL-SCCA ensures that correlated features are grouped together and given similar canonical weights, whether they are positively or negatively correlated, which gives the GGL-SCCA an advantage over previous SS-CCA methods. On an example of two features from one view, the grouping effect follows:

$$|u_i^* + u_j^*| \leq \frac{(1+\alpha_1)\sqrt{u_i^{*2}+u_j^{*2}}}{\lambda_1}\sqrt{2(1+\rho_{ij})} \qquad (4)$$

Where u^* is the solution to the SS-CCA problem. u_i and u_j are features connected in the graph. ρ_{ij} denotes sample correlation between connected features.

Pairing-based structured ensemble classifier learning. Next, for each pair of multiplexes, we concatenate highly correlated features sparsely selected from each multiplex to train a linear support vector machine (SVM) classifier (**Fig. 3**). Noting that for each training subject we have N estimated multiplexes, we perform C_N^2 mappings of each pair of multiplexes in $\mathbb{M}$. Subsequently, a linear SVM classifier is learned for each pair of multiplexes. In the testing stage, for a specific pair of multiplexes, we use the weights **u** and **v** learned for this pair to respectively select features from each testing multiplex, which are then inputted to the pair-specific trained SVM classifier. Finally, we combine the decisions from all individual SVM classifiers in the ensemble through a weighted voting strategy (i.e. by averaging all prediction scores) to predict the label of the testing subject.

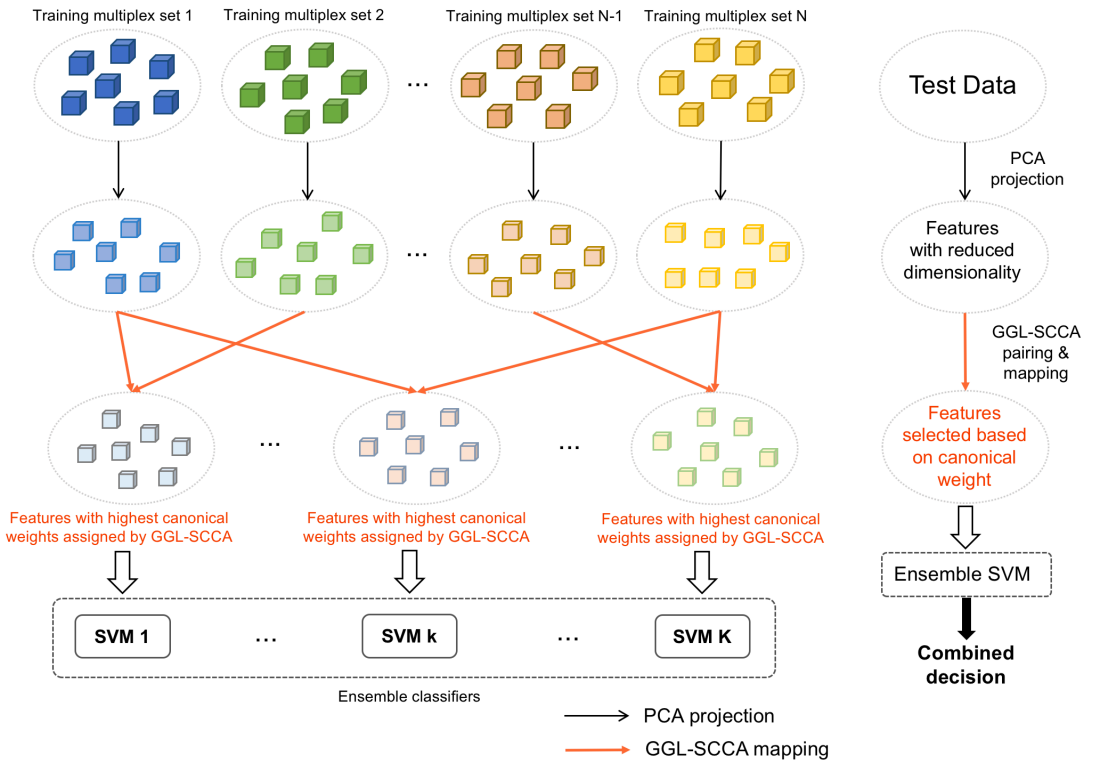


Figure 3: *Pipeline of the proposed pairing-based structured ensemble classifier learning using fused shallow convolutional brain multiplexes*. We use principal component analysis (PCA) to first reduce the dimensionality of each training multiplex set, then apply structure-constrained sparse canonical correlation analysis (GGL-SCCA) (Du et al., 2017) to map a pair of PCA-projected multiplex sets onto a common space where pairs of multiplexes are most correlated. Then we fuse the selected highly correlated features (given highest canonical weights) from *paired* shallow multiplexes to train a linear SVM classifier. The decisions from all the individual classifiers in the ensemble are combined for the final classification decision.

Identification of highly correlated morphological connections. Since we need to use PCA for multiplex dimensionality reduction prior to GGL-SCCA mapping to avoid the covariance matrix singularity, we are not able to directly identify the selected multiplex morphological features. However, this projection step can be avoided when using pairs of views, as they have a much smaller feature vector size than that of multiplexes. Hence, for morphological connectional features identification, we performed the GGL-SCCA mapping directly on the pairs of views without previous application of any dimensionality reduction technique. To do so, for each GGL-SCCA mapped morphological view, the canonical weights of each feature were averaged

across subjects and when paired with other views. Ultimately, top K features with highest average canonical weights were identified. GGL-SCCA is first performed on the training data. The features were given canonical weights. Based on these weights, the most relevant features were identified. For each feature, we take the average of its canonical weights obtained across different subjects to identify features with the highest average weight.

2.2 Materials

Evaluation Dataset. We used 42 eMCI (average age 70.4 ± 7.5) and 42 NC (average age 74.1 ± 6.7) age and gender-matched subjects from ADNI GO public dataset[§], each with structural T1-w MR image (Mueller *et al.*, 2005). We used FreeSurfer analysis suite[**] to reconstruct both right and left cortical hemispheres for each subject from T1-w MRI. The processing included skull stripping, motion correction, two T1-w images averaging, intensity normalization, topology correction and segmentation of the subcortical White Matter (WM) and deep Grey Matter (GM) volumetric structures to identify GM/WM and GM/Cerebrospinal fluid (CSF) boundaries, as in (Dale *et al.*, 1999). Then we parcellated each cortical hemisphere into 35 cortical regions using Desikan-Killiany atlas.

Proposed brain multiplexes. We defined $N = 6$ shallow multiplexes, each using 2 cortical network views. For each cortical attribute, we compute the strength of the morphological network connection linking ith ROI to the jth ROI as the absolute difference between the averaged attribute values in both ROIs. Multiplex $\boldsymbol{M_1}$ includes cortical attribute views $\{\boldsymbol{V_1}, \boldsymbol{V_2}\}$, $\boldsymbol{M_2}$ includes $\{\boldsymbol{V_1}, \boldsymbol{V_3}\}$, $\boldsymbol{M_3}$ includes $\{\boldsymbol{V_1}, \boldsymbol{V_4}\}$, $\boldsymbol{M_4}$

[§] http://adni.loni.usc.edu

[**] https://surfer.nmr.mgh.harvard.edu

includes $\{\boldsymbol{V_2}, \boldsymbol{V_3}\}$, $\boldsymbol{M_5}$ includes $\{\boldsymbol{V_2}, \boldsymbol{V_4}\}$, and $\boldsymbol{M_6}$ includes $\{\boldsymbol{V_3}, \boldsymbol{V_4}\}$. For each cortical region, $\boldsymbol{V_1}$ denotes the maximum principal curvature brain view, $\boldsymbol{V_2}$ denotes the mean cortical thickness brain view, $\boldsymbol{V_3}$ denotes the mean sulcal depth brain view, and $\boldsymbol{V_4}$ denotes the mean of average curvature brain view.

Remark: The morphological networks and multiplexes are constructed separately for the left and the right hemispheres and they are studied independently, as we do not want our morphological connections to be 'biased' by brain hemispheric asymmetry. It also prevents loss of insightful information on how eMCI affects each hemisphere *independently*.

3. Results

Experimental setup. We used leave-one-out cross validation strategy with SVM parameter C (strength of penalty imposed on miss-classified data) set at default $\mathtt{C} = 1$, which aids the training efficiently, while avoiding overfitting of the model to the training data. Since the grouping effect of GGL-SCCA is controlled oppositely by $\alpha_{1,2}$ and $\lambda_{1,2}$ and it is more sensitive to $\lambda_{1,2}$ than to $\alpha_{1,2}$, we fix $\alpha_{1,2}$ at a moderate value; $\alpha_1 = \alpha_2 = 10$, as indicated in (Du *et al.*, 2017). Then, we fine tuned λ_1 and λ_2 using 5-fold nested cross-validation on the training data. If $\lambda_{1,2}$ is too small, the GGL-SCCA would reduce to CCA and too large $\lambda_{1,2}$ leads to over-penalization of the results. Therefore, we set the range for the grid search of $\lambda_{1,2}$ between 0.1 and 0.5, which gives the desired number of features and accuracy level for the classification. Exhaustive hyper-parameter tuning is not feasible due to high computational power requirements. However, our experiments show that additional tuning of C and α, do not significantly

affect the performance. The range for λ tuning was established experimentally for left and right hemispheres.

For the classification task using paired GGL-SCCA mapping, due to very high dimensionality of multiplex structures, the pairs of multiplexes were first projected onto a lower dimensional space using PCA. For a fair comparison of performance, we also projected the pairs of views onto a lower dimensional space. According to (Park *et al.*, 2012), the PCA might aid in improving the classification performance of early demented patients from healthy controls using morphological brain features. Since, the performance of classifiers heavily depends on the number of input features, for the SS-CCA we chose a range of input features given the highest canonical weights. The ensemble SVM classification performance was evaluated using the top $\{10, 15, \ldots, 45, 50\}$ features selected from each view/multiplex in the pair. For the final SS-CCA classification performance, we report the average classification accuracy for different feature numbers in $\{10, 20, \ldots, 50\}$.

Comparison methods. To the best of our knowledge, no other study used morphological brain networks for eMCI diagnosis. Therefore, we benchmark our proposed framework against other similar approaches using our morphological data. For our eMCI/NC classification task, we benchmarked our pairing-based ensemble classifier strategy against: (1) using single SVM trained on the concatenated views, (2) ensemble SVM classifiers (without the pairing or any mapping strategies), (3) ensemble paired SVM classifiers (without CCA mapping) and (4) ensemble paired SVM classifiers with CCA mappings (Lisowska and Rekik, 2017). For each of these methods, we generated four classification results using: (1) features from brain views, (2) features from correlational 4-layer multiplexes (inter-layer computed using Pearson correlation), (3) features from convolutional 4-layer multiplexes (interlayer computed

using 2D convolution) and (4) features from shallow (2-layer) convolutional multiplexes. For evaluation, we report in **Table 2** the classification accuracy, the area under the receiver operating characteristic (ROC) curve, the sensitivity and specificity of the eMCI/NC classification task. In **Fig. 4**, we specifically show the comparison of classification accuracy for ensemble, pairing-based ensemble (without CCA or GGL-SCCA mapping), pairing-based ensemble with CCA-mapping and pairing-based ensemble with GGL-SCCA mapping classification based on (1) concatenated views, (2) correlational deep multiplexes, (3) convolutional deep multiplexes (Lisowska and Rekik, 2017), and (4) proposed shallow convolutional multiplexes.

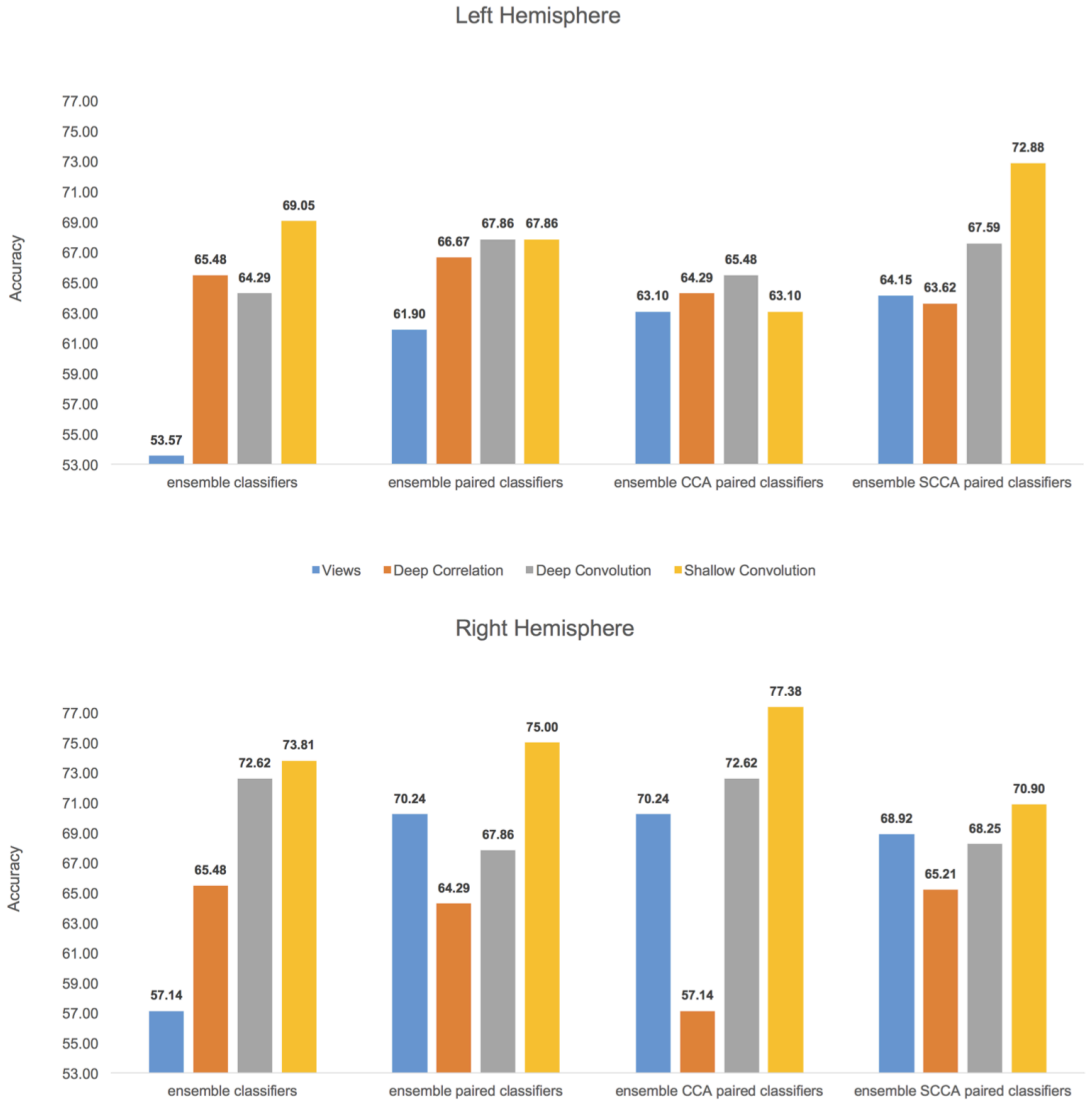


Figure 4: *Classification accuracies for our proposed joint pairing and structured GGL-SCCA mapping of brain features and comparison with other ensemble classifier methods.* Views: morphological brain views. Correlation: correlational brain multiplexes. Convolution: Convolutional brain multiplexes. 2-layer: shallow convolutional multiplex. Ensemble classifiers: one SVM trained for each view (or multiplex) without any pairing strategy or mapping. Ensemble paired classifiers: pairing different views (or multiplexes) without any mapping. Ensemble CCA paired classifiers: pairing different views (or multiplexes) with CCA mapping, Ensemble GGL-SCCA paired classifiers: pairing different views (or multiplexes) with GGL-SCCA mapping (average accuracy across different number of features from 10 to 50, with an incremental step of 10 features). For brevity, we shortened GGL-SCCA to SCCA.

Table 2. *eMCI/NC classification performance using our method and different comparison methods.*

		Left Hemisphere				Right Hemisphere			
Classifier	Method	Accuracy (%)	AUC	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUC	Sensitivity (%)	Specificity (%)
ensemble classifiers	Views	53.57	54.31	57.14	50.00	57.14	62.24	52.38	61.90
	Deep Correlation	65.48	67.18	66.67	64.29	65.48	70.52	66.67	64.29
	Deep Convolution	64.29	73.81	69.05	59.52	72.62	74.21	71.43	73.81
	Shallow Convolution	69.05	72.85	**73.81**	64.29	73.81	76.59	73.81	73.81
ensemble paired classifiers	Views	61.90	73.58	64.29	59.52	70.24	73.53	71.43	69.05
	Deep Correlation	66.67	67.86	69.05	64.29	64.29	70.52	66.67	61.90
	Deep Convolution	67.86	72.68	71.43	64.29	67.86	71.20	66.67	69.05
	Shallow Convolution	67.86	73.07	71.43	64.29	75.00	75.96	73.81	**76.19**
ensemble CCA paired classifiers	Views	63.10	66.89	66.67	59.52	70.24	76.87	69.05	71.43
	Deep Correlation	64.29	67.01	69.05	59.52	57.14	63.21	64.29	50.00
	Deep Convolution	65.48	64.40	66.67	64.29	72.62	78.40	76.19	69.05
	Shallow Convolution	63.10	67.06	64.29	61.90	**77.38**	**79.20**	**78.57**	**76.19**
ensemble GGL-SCCA paired classifier	Views	64.15	72.06	61.11	67.20	68.92	73.50	69.58	68.25
	Deep Correlation	63.62	67.11	61.11	66.14	65.21	70.67	68.25	62.17
	Deep Convolution	67.59	74.29	65.08	70.11	68.25	70.93	69.58	66.93
	Shallow Convolution	**72.88**	**76.06**	66.93	**78.84**	70.90	76.52	70.11	71.69

The best accuracy, area under the ROC curve, sensitivity and specificity were always obtained when using the proposed shallow convolutional multiplexes

(**Table 2**). Our proposed ensemble GGL-SCCA paired classifier framework outperformed all comparison methods (72.9%) when using shallow convolutional multiplexes of the left hemisphere (**Fig. 4** and **Table 2**). On the other hand, the best classification result (77.4%) for the right hemisphere was obtained using ensemble CCA paired classifier, introduced in (Lisowska and Rekik, 2017). This might indicate that the group effect introduced by GGL-SCCA may better capture a morphological connectional structure in the left hemisphere that is not prevalent in the right hemisphere. It might also imply different levels of complexity in morphological cortical disease progression across both hemispheres.

Identification of highly correlated morphological connections. For each morphological view, we identified the top 10 features with the highest average GGL-SCCA canonical weights across subjects and when paired with other views for both left and right hemispheres. **Fig. 5** displays circular graphs with top 10 features identified by GGL-SCCA for each cortical attribute and each hemisphere. We display in **Fig. 6** the index of each anatomical cortical brain region referred to in the circular graph together with their corresponding names.

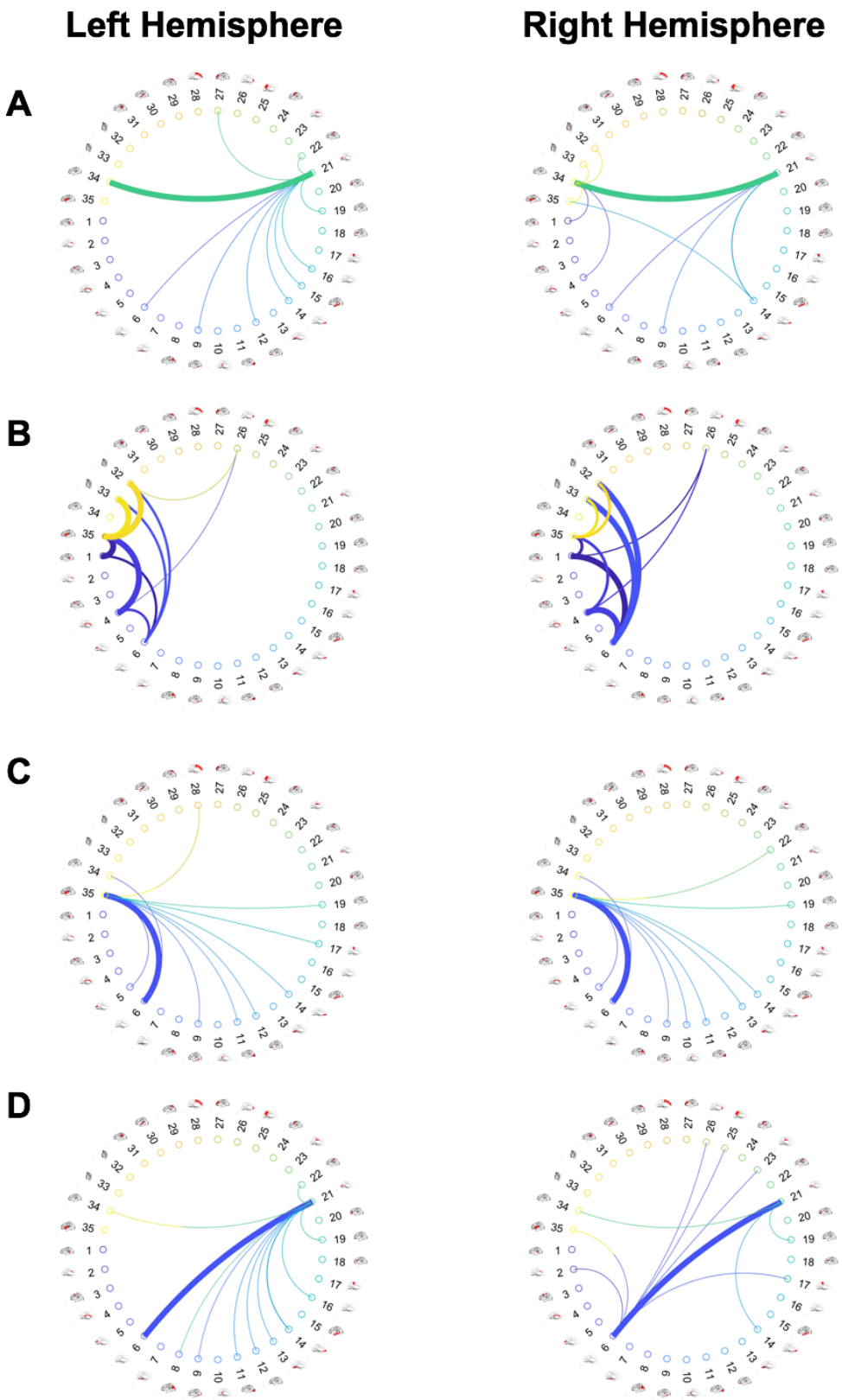
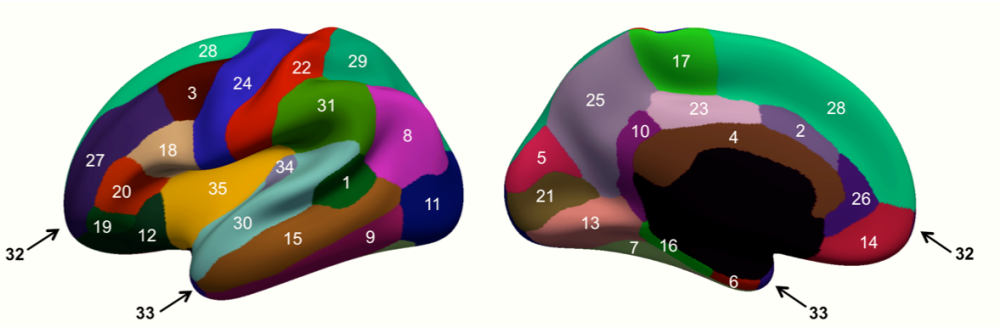


Figure 5: *The top 10 connections given highest canonical weights by the* GGL-SCCA *visualized on the circular graphs for the left and the right hemisphere*. A) The maximum principal curvature brain view. B) The mean cortical thickness brain view. C) The mean sulcal depth brain view. D) The mean of average curvature.



1. Bank of the Superior Temporal Sulcus	**2.** Caudal Anterior-cingulate Cortex
3. Caudal Middle Frontal Gyrus	**4.** Unmeasured Corpus Callosum
5. Cuneus Cortex	**6.** Entorhinal Cortex
7. Fusiform Gyrus	**8.** Inferior Parietal Cortex
9. Inferior Temporal Gyrus	**10.** Isthmus-cingulate Cortex
11. Lateral occipital cortex	**12.** Lateral orbital frontal cortex
13. Lingual gyrus	**14.** Medial orbital frontal cortex
15. Middle temporal gyrus	**16.** Parahippocampal gyrus
17. Paracentral lobule	**18.** Pars opercularis
19. Pars orbitalis	**20.** Pars triangularis
21. Pericalcarine cortex	**22.** Postcentral gyrus
23. Posterior-cingulate cortex	**24.** Precentral gyrus
25. Precuneus cortex	**26.** Rostral anterior cingulate cortex
27. Rostral middle frontal gyrus	**28.** Superior frontal gyrus
29. Superior parietal cortex	**30.** Superior temporal gyrus
31. Supramarginal gyrus	**32.** Frontal pole
33. Temporal pole	**34.** Transverse temporal cortex
35. Insula cortex	

Figure 6: *Cortical regions of interest index. Each cortical hemisphere is parcellated using Desikan Cortical atlas*. We display the cortical region names and their respective identification numbers.

For the three cortical attributes (the maximum principal curvature, the mean sulcal depth and the mean of average curvature), one very strongly correlated connection was present (**Fig. 5-A,C,D**), whereas for mean cortical thickness, more highly correlated connections were identified (**Fig. 5-B**). The connection given the highest canonical weight in the maximum principal curvature brain view was the connection between the transverse temporal cortex and the pericalcarine cortex, which seems to act as a 'hub' region for many highly correlated connections. This is consistent for both cortical hemispheres. Additionally, the transverse temporal cortex seems to serve as a hub

region in the right hemisphere (**Fig. 5-A**). For the mean sulcal depth brain view (**Fig. 5-C**), the strongest connection was between the entorhinal cortex and the insula cortex, which is identified as a hub region with many connections extending from it in the left and right hemisphere. On the other hand, for the mean of average curvature brain view (**Fig. 5-D**), we note a distinctive strong connection between the entorhinal cortex and the pericalcarine cortex, which acts as a hub for other morphological connections in the left hemisphere, as well as the right hemisphere. For the mean cortical thickness brain view (**Fig. 5-B**), many highly correlated connections were spread out across the frontal pole, the temporal pole, the insula cortex, the bank of superior temporal sulcus, the corpus callosum and the entorhinal cortex.

Having identified the underlying structure of important connections between brain regions, we wanted to identify which connections are altered by eMCI and could serve as early disease marker. To do so, we plotted a graph of classification accuracies obtained from each GGL-SCCA mapped pair of views in the ensemble (**Fig. 7**) to identify the most discriminative features between NC and eMCI patients. These are obtained from pairs of views that gave the highest classification accuracy.

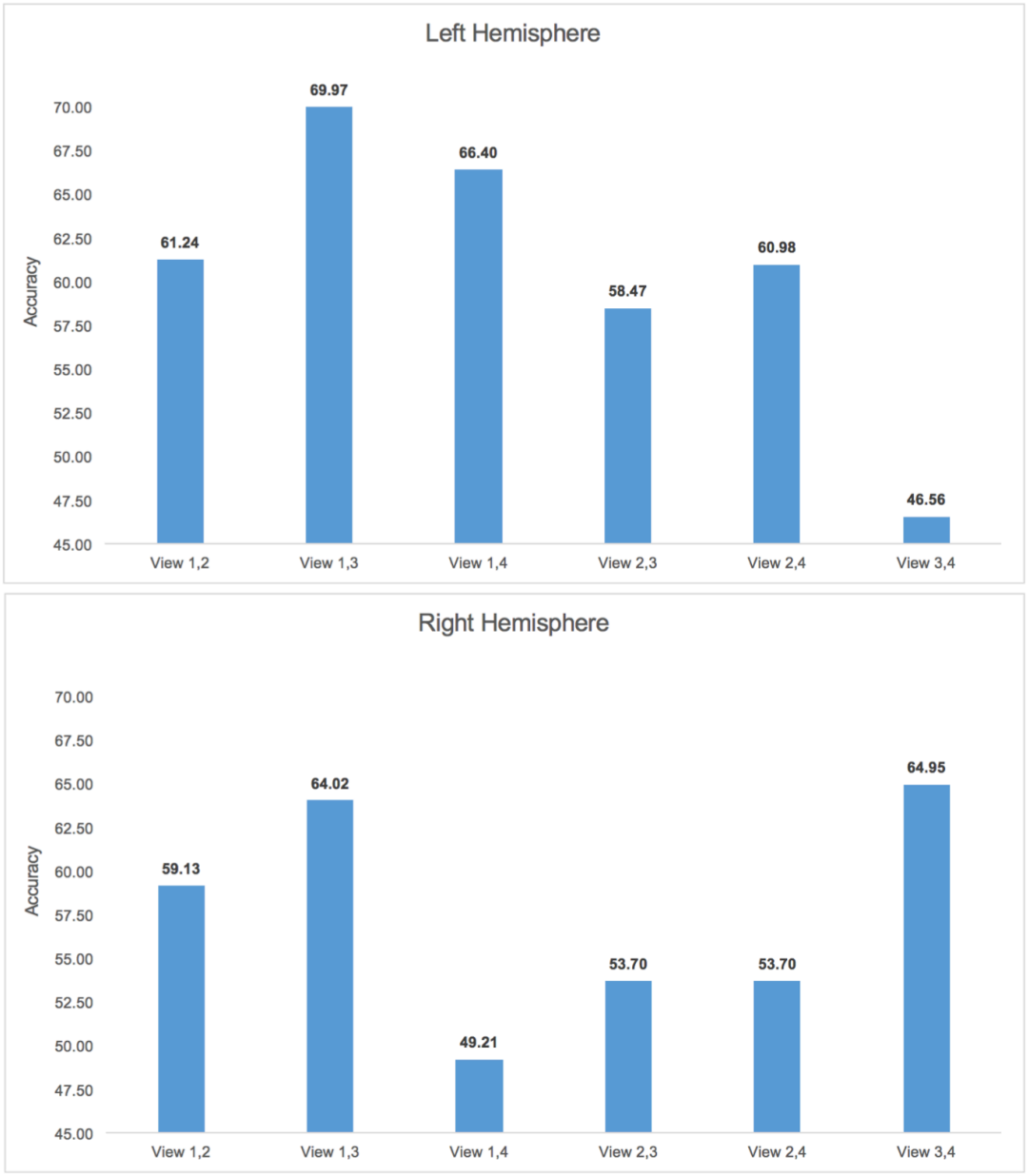


Figure 7: *Classification accuracy obtained from different pairs of morphological views in the ensemble paired* GGL-SCCA *mapping setting*. 1: the maximum principal curvature brain view. 2: the mean cortical thickness brain view. 3: the mean sulcal depth brain view. 4: the mean of average curvature.

Influence of PCA on classification performance. We also compared the performance of ensemble paired views with GGL-SCCA mapping with and without prior PCA (**Fig. 7**). The PCA-projected morphological views gave higher classification accuracy for the left and the right hemispheres as compared to using raw morphological

views across all the different number of features used for classification using the ensemble GGL-SCCA classifier.

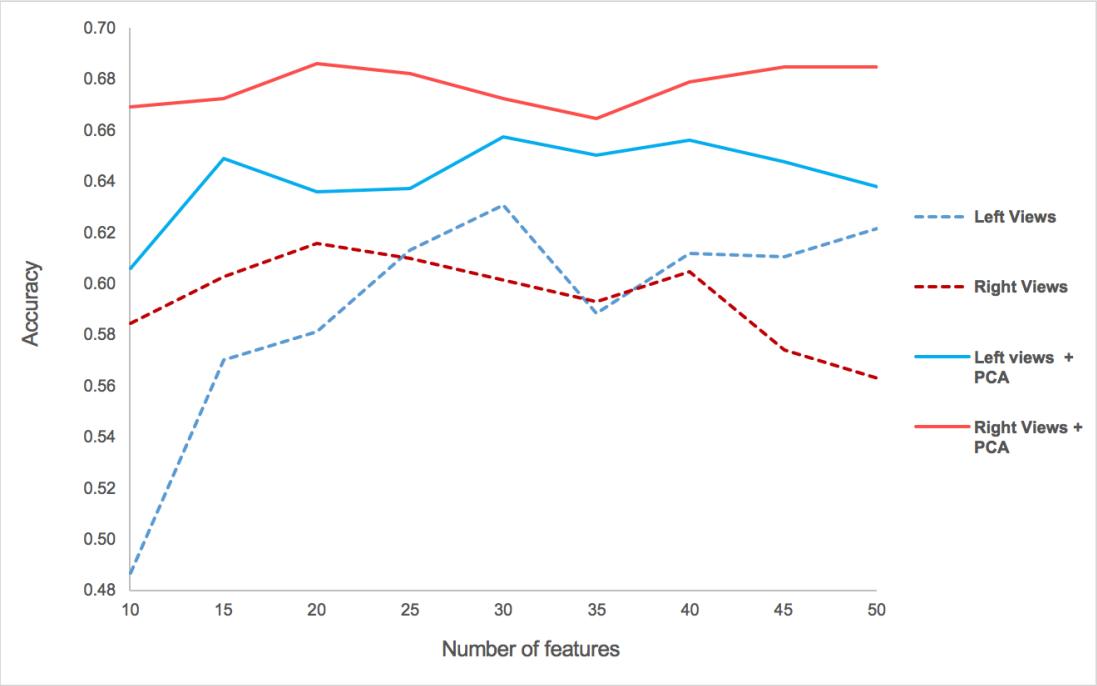


Figure 8: *Classification accuracies using the pairing-based ensemble classifier learning using GGL-SCCA-mapped brain view features with and without dimensionality reduction*. Left views: morphological brain views from the left hemisphere. Right views: morphological brain views from the right hemisphere. Left views + PCA: PCA-projected brain views from the left hemisphere. Right views + PCA: PCA-projected brain views from the right hemisphere.

Ensemble classifier learning strategy for early MCI diagnosis. The ensemble GGL-SCCA paired classifier performed better than using any single combination of multiplexes when using the shallow convolutional brain multiplexes as input features (**Fig. 9**), indicating that ensemble classifier is an effective strategy in combining information from different morphological brain properties to improve early dementia diagnosis.

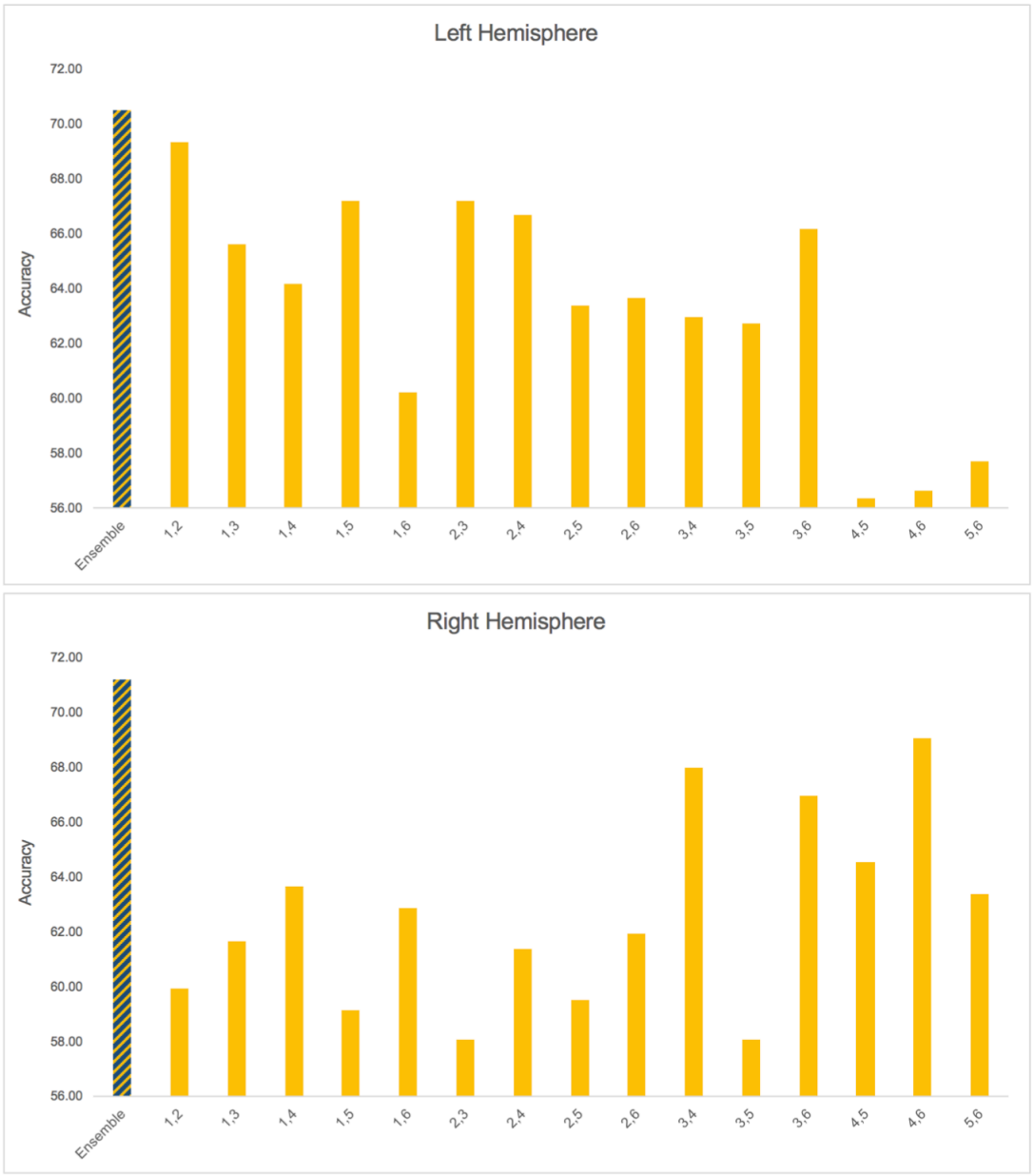


Figure 9: *Classification accuracy using different pairings of our six shallow (2-layer) convolutional multiplexes in the ensemble paired* GGL-SCCA *mapping setting*. Numbers displayed along the horizontal axis denote the multiplex index (e.g., 1 for $\boldsymbol{M_1}$). The multiplex pair $(\boldsymbol{M_1}, \boldsymbol{M_2})$ of the left hemisphere produced the best classification accuracy among individual networks, while multiplex $(\boldsymbol{M_4}, \boldsymbol{M_6})$ of the right hemisphere produced the best classification accuracy. For both hemispheres, ensemble multiplexes outperformed individual multiplexes.

4. Discussion

We proposed a joint pairing and structured mapping strategy of morphological brain multiplexes for early dementia diagnosis. Specifically, we presented an individual-specific representation of brain connectivity based on shallow convolutional brain multiplexes, each capturing a *unique relationship* between two different cortical networks. To combine the information from these various multiplexes, we introduced a pairing-based ensemble classifier, where we fused features sparsely selected from a pair of multiplexes using GGL-SCCA mapping, trained a linear SVM classifier for each of the mapped features, and combined the decisions from all the classifiers in the ensemble by weighted voting to distinguish eMCI patients from NC subjects.

Our proposed ensemble GGL-SCCA paired classifier outperformed other methods when using shallow convolutional multiplexes based on morphological brain features from the left hemisphere, while the best classification result for the right hemisphere was obtained using CCA-mapped shallow convolutional multiplexes (**Fig. 4** and **Table 2**). These results may suggest that early dementia exhibits different morphological progression patterns across the two cortical hemispheres, possibly affecting the main communication centers (hubs) predominantly in the left hemisphere, while affecting the right hemisphere in a less structured manner. It also shows that for the right hemisphere, the correlation between different morphological attributes was most discriminative between the eMCI and NC patient, while the structure information introduced by SS-CCA may be more prevalent in the left hemisphere.

PCA dimensionality reduction. We chose to use PCA to project our data onto a lower dimensional space, since it was shown to improve the discriminative power of brain morphological features in classifying MCI and NC subjects (Park *et al.*, 2012). In our work, we compared the classification accuracy of the ensemble GGL-SCCA paired classifier using morphological brain views with and without previous application of

PCA and confirmed that reducing the dimensionality of morphological data using PCA improves the eMCI/NC classification performance (**Fig. 7**). However, PCA as an unsupervised method, does not take into account the data class labels and it does not ensure good separability of the data from the two different classes in the new space. Furthermore, projecting the morphological features onto a lower dimensional space hinders their biological interpretability, making it impossible to identify important multiplex biomarkers for early dementia diagnosis.

Identification of highly correlated morphological features. Despite these drawbacks and thanks to the sparsity constraint of the GGL-SCCA, we were able to apply GGL-SCCA to pairs of morphological brain networks and identify features that are highly correlated between different cortical attributes. We found that the most discriminative features for the eMCI diagnosis belonged to maximum principal curvature and average curvature views (**Fig. 7**). These produced the highest classification accuracies when paired with sulcal depth brain view, thereby indicating correlated brain changes in sulcal depth and curvature in early dementia. Further, GGL-SCCA allowed us to identify underlying structure of the brain networks for different cortical attributes. Specifically, hub regions, which were characterized by highly correlated morphological connections to many other cortical regions, were identified in the entorhinal cortex and the insula cortex (**Fig. 5**). These regions and their morphological relationships may serve as important biomarkers for early MCI diagnosis. The insula cortex is involved in emotion processing and regulation, introspective awareness and integration of multimodal inputs and the reduction in the insular volume was associated with the occurrence of neuropsychiatric symptoms, such as agitation and apathy in patients with Alzheimer's disease (Moon *et al.*, 2014). The entorhinal cortex plays an important role in navigation and formation and consolidation of spatial and declarative memory.

Various research works, in line with our findings, have shown that the pathological changes in the neuronal structure and function of the entorhinal cortex occur before the onset of any AD symptoms (Zhou *et al.*, 2016), making it a region of interest for early dementia diagnosis.

To get further insights into which of the identified features are most discriminative between the NC and eMCI patients, we plotted a graph of classification accuracies obtained from each GGL-SCCA mapped pair of views in the ensemble (**Fig. 8**). The highest accuracy of NC/eMCI classification (69.97%) using cortical attributes from the left hemisphere was obtained by a classifier trained on a pair of views consisting of the maximum principal curvature and the mean sulcal depth. This might indicate that the connection between the transverse temporal cortex and the pericalcarine cortex and the connection between the entorhinal cortex and the insula cortex may be a likely target of morphological changes in the early stages of dementia. Interestingly, the pairs which contained the maximum principal curvature as one of the views, consistently performed best in case of the left hemisphere, which shows that this cortical attribute fingerprints 'morphological' dementia progression. For the right hemisphere, the best classification performance was also achieved when using the sulcal depth and the mean of average curvature, giving 64.95% accuracy (**Fig. 8**). The most notable connections are between the entorhinal cortex and the pericalcarine cortex and the entorhinal cortex and the insula cortex. Together, these results suggest that the curvature of the sulci is also altered by the early dementia and could provide insight into the progression of the disease. Further, the most likely candidates of hub regions affected by the eMCI are the entorhinal cortex and the pericalcarine cortex.

Our findings are consistent with previous research on MCI patients, which reported brain atrophy in regions such as the entorhinal cortex, hippocampus, medial temporal

lobe, insula cortex and temporal lobe (Fan *et al.*, 2008). Additionally, (Li *et al.*, 2016) showed that networks constructed by sulcal depth exhibit disrupted connectivity in amnestic MCI patients, implying loss of efficiency in communication between different brain regions. Recently, (Cai *et al.*, 2017), which used morphological properties for MCI diagnosis, showed that the most discriminative features were found in the hippocampus, amygdala, entorhinal cortex and precuneus. However, none of the previous research explored the *relationship* between *morphological networks* constructed from different cortical attributes as in our current study.

Shallow vs deep multiplex structure. Our research showed that introducing the relationship between different cortical networks (convolutional interlayers exploring the relationship between brain views), which we modeled using a multiplex structure, improved the classification accuracy of eMCI/NC in comparison to using different brain networks alone (**Fig. 9**). The new 2-layer multiplex structure reduced the correlation between individual classifiers in the ensemble and resulted in a better ensemble classifier performance comparing to the ensemble classifier using previous convolutional multiplex structure (**Fig. 4** and **Table 2**). Furthermore, we show that the ensemble classifier combining prediction scores from linear SVM classifiers using shallow multiplexes better distinguishes eMCI patients from NC subjects than any of the individual classifiers used independently for both, left and right hemisphere (**Fig. 9**).

Limitations and future directions. With increasing availability of data and computational power, more advanced machine learning-based classifiers, e.g., structured random forests and neural networks, can be used as base classifiers to define our ensemble to improve early dementia diagnosis. However, one advantage of using linear classifiers as individual classifiers in the ensemble, is that they generally give

stable results —i.e., the output from the ensemble does not significantly change in response to small changes in the training data (Dietterich *et al.*, 2000).

Although we showed that the relationship between different cortical brain networks brings novel insights into early dementia and cortex morphology from multiple views, we were unable to pinpoint the specific relationship between brain connections from different morphological brain networks that are most discriminative between the eMCI and NC classes due to the necessity of applying previous dimensionality reduction method before the GGL-SCCA mapping. In fact, our method cannot directly identify the informative connectional features, however, it achieves the best classification accuracy. For this reason, we used the paired views to identify the most discriminative connectional features. We also note the devised architecture of pairing morphological brain views in our ensemble learning framework is novel. Ideally, our multiplex-based classification method would be able to do achieve both tasks (high classification accuracy and identifying discriminative features). However, this was not possible due to the curse of using projection for feature dimensionality reduction to well-train a classifier, which hinders the tracking of the original features. This is a typical drawback of classification methods that make use of projections such as conventional CCA and Linear Discriminant Analysis (LDA). Future work should seek a mapping method that would be less sensitive to the high dimensionality of the data or a dimensionality reduction technique (i.e. feature selection) that would not adversely hinder the interpretability of the projected multiplex features.

In addition to investigating how dementia alters the relationship between connections from different cortical networks, we expect that by jointly exploring changes in the brain morphological, structural and functional networks, we could provide a more *holistic connectomic understanding* of early dementia onset and further

improve the classification performance of our model for a more reliable early dementia diagnosis. If this study is further extended and improved to match the performance of eMCI diagnosis based on structural/functional connectivity, the time and cost of eMCI diagnosis would be significantly reduced and more patients could benefit from early dementia diagnosis.

5. Conclusion

In this work, we proposed a joint pairing and mapping strategy for early dementia diagnosis using morphological convolutional brain multiplexes. This allowed to investigate multi-view changes in brain cortical morphology on a connectional level. Our pairing-based ensemble classifier learning strategy produced the best classification performance in comparison to state-of-the-art methods. More importantly, we identified the most discriminative pairs of morphological brain views distinguishing eMCI from NC subjects as well as the most highly correlated morphological brain connections in our cohort. In our future work, we will focus on further improving the data fusion method (i.e. by incorporating supervised multi-view mapping or applying more complex learning models) and integrating multimodal connectomic data (e.g., functional and structural networks) to complement our morphological connectivity data for more holistic investigation of early dementia.

Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Araclon

Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Author disclosure statement.

No competing financial interests exist.

References

Brown, C.J., Hamarneh, G., 2016. Machine learning on human connectome data from mri. arXiv preprint arXiv:1611.08699 .

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nature reviews. Neuroscience 10, 186.

Cai, K., Xu, H., Guan, H., Zhu, W., Jiang, J., Cui, Y., Zhang, J., Liu, T., Wen, W., 2017. Identification of early-stage alzheimer's disease using sulcal morphology and other common neuroimaging indices. PloS one 12, e0170875.

Chen, J., Bushman, F.D., Lewis, J.D., Wu, G.D., Li, H., 2013. Structureconstrained sparse canonical correlation analysis with an application to microbiome data analysis. Biostatistics 14, 244–258.

Chen, X., Liu, H., 2012. An efficient optimization algorithm for structured sparse cca, with applications to eqtl mapping. Statistics in Biosciences 4, 3–26.

Chen, X., Zhang, H., Gao, Y., Wee, C.Y., Li, G., Shen, D., 2016. High-order resting-state functional connectivity network for mci classification. Human brain mapping 37, 3282–3296.

Crofts, J., Forrester, M., O'Dea, R., 2016. Structure-function clustering in multiplex brain networks. EPL (Europhysics Letters) 116, 18003.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis: I. segmentation and surface reconstruction. Neuroimage 9, 179–194.

Dietterich, T.G., *et al.*, 2000. Ensemble methods in machine learning. Multiple classifier systems 1857, 1–15.

Du, L., Huang, H., Yan, J., Kim, S., Risacher, S.L., Inlow, M., Moore, J.H., Saykin, A.J., Shen, L., Initiative, A.D.N., 2016. Structured sparse canonical correlation analysis for brain imaging genetics: an improved graphnet method. Bioinformatics 32, 1544–1551.

Du, L., Zhang, T., Liu, K., Yan, J., Yao, X., Risacher, S.L., Saykin, A.J., Han, J., Guo, L., Shen, L., *et al.*, 2017. Identifying associations between brain imaging phenotypes and genetic factors via a novel structured SCCA approach , 543–555. Džeroski, S., Ženko, B., 2004. Is combining classifiers with stacking better than selecting the best one? Machine learning 54, 255–273.

Fan, Y., Batmanghelich, N., Clark, C.M., Davatzikos, C., Initiative, A.D.N., *et al.*, 2008. Spatial patterns of brain atrophy in mci patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. Neuroimage 39, 1731–1743.

Giuliano Zippo, A., Castiglioni, I., 2016. Integration of 18fdg-pet metabolic and functional connectomes in the early diagnosis and prognosis of the alzheimer's disease. Current Alzheimer Research 13, 487–497.

Haghighat, M., Abdel-Mottaleb, M., Alhalabi, W., 2016. Fully automatic face normalization and single sample face recognition in unconstrained environments. Expert Systems with Applications 47, 23–34.

Hamelin, L., de Souza Leonardo, C., Corlier, F., Corne, H., Chupin, M., Dubois, B., Bottlaender, M., Colliot, O., Sarazin, M., 2014. Improved accuracy of the diagnosis of

early Alzheimers disease using combined measures of hippocampal volume and sulcal morphology (p4. 016). Neurology 82, P4–016.

Im, K., Lee, J.M., Seo, S.W., Kim, S.H., Kim, S.I., Na, D.L., 2008. Sulcal morphology changes and their relationship with cortical thickness and gyral white matter volume in mild cognitive impairment and alzheimer's disease. Neuroimage 43, 103–113.

La Rocca, M., Amoroso, N., Bellotti, R., Diacono, D., Monaco, A., Monda, A., Tateo, A., Tangaro, S., 2017. A multiplex network model to characterize brain atrophy in structural mri, in: Emergent Complexity from Nonlinearity, in Physics, Engineering and the Life Sciences. Springer, pp. 189–198.

Li, Q., Li, X., Wang, X., Li, Y., Li, K., Yu, Y., Yin, C., Li, S., Han, Y., 2016. Topological properties of large-scale cortical networks based on multiple morphological features in amnestic mild cognitive impairment. Neural plasticity 2016.

Lisowska, A., Rekik, I., and and The Alzheimer's Disease Neuroimaging Initiative, 2017. Pairing-based ensemble classifier learning using convolutional brain multiplexes and multi-view brain networks for early dementia diagnosis, In International Workshop on Connectomics in Neuroimaging, 42–50.

Liu, T., Lipnicki, D.M., Zhu, W., Tao, D., Zhang, C., Cui, Y., Jin, J.S., Sachdev, P.S., Wen, W., 2012. Cortical gyrification and sulcal spans in early stage alzheimer's disease. PLoS One 7, e31083.

Mahjoub, I., Mahjoub, M.A. and Rekik, I., 2018. Brain multiplexes reveal morphological connectional biomarkers fingerprinting late brain dementia states. *Scientific reports*, *8*(1), p.4103.

Moon, Y., Moon, W.J., Kim, H., Han, S.H., 2014. Regional atrophy of the insular cortex is associated with neuropsychiatric symptoms in alzheimer's disease patients. European neurology 71, 223–229.

Mucke, L., 2009. Neuroscience: Alzheimer's disease. Nature 461, 895–897.

Mueller, S.G., Weiner, M.W., Thal, L.J., Petersen, R.C., Jack, C., Jagust, W., Trojanowski, J.Q., Toga, A.W., Beckett, L., 2005. The Alzheimer's Disease Neuroimaging Initiative. Neuroimaging Clinics of North America 10, 869–877.

Park, H., Yang, J.j., Seo, J., Lee, J.m., 2012. Dimensionality reduced cortical features and their use in the classification of alzheimer's disease and mild cognitive impairment. Neuroscience letters 529, 123–127.

Parkhomenko, E., Tritchler, D., Beyene, J., 2009. Sparse canonical correlation analysis with application to genomic data integration. Statistical applications in genetics and molecular biology 8, 1–34.

Prasad, G., Joshi, S.H., Nir, T.M., Toga, A.W., Thompson, P.M., (ADNI, A.D.N.I., *et al.*, 2015. Brain connectivity and novel network measures for alzheimer's disease classification. Neurobiology of aging 36, S121–S131.

Prince, M., Prina, M., Guerchet, M., 2013. Journey of Caring: an analysis of long-term care for Dementia. Ph.D. thesis. N/A Ed; London: Alzheimer's Disease International.

Quan, Y., Xu, Y., Sun, Y., Huang, Y., Ji, H., 2016. Sparse coding for classification via discrimination ensemble , 5839–5847.

Querbes, O., Aubry, F., Pariente, J., Lotterie, J.A., D´emonet, J.F., Duret, V., Puel, M., Berry, I., Fort, J.C., Celsis, P., *et al.*, 2009. Early diagnosis of alzheimer's disease using cortical thickness: impact of cognitive reserve. Brain 132, 2036–2047.

Soussia, M. and Rekik, I., 2017. High-order Connectomic Manifold Learning for Autistic Brain State Identification. In *International Workshop on Connectomics in Neuroimaging*, 51-59.

Singanamallia, A., Wang, H., Lee, G., Shih, N., Rosen, M., Master, S., Tomasewski, J., Feldman, M., Madabhushi, A., 2014. Supervised multiview canonical correlation analysis: Fused multimodal prediction of disease diagnosis and prognosis 9038, 903805–1.

Van Essen, D.C., 1997. A tension-based theory of morphogenesis and compact wiring in the central nervous system. Nature 385, 313.

Wee, C.Y., Yang, S., Yap, P.T., Shen, D., Initiative, A.D.N., *et al.*, 2016. Sparse temporally dynamic resting-state functional connectivity networks for early mci identification. Brain imaging and behavior 10, 342–356.

Yao, Z., Zhang, Y., Lin, L., Zhou, Y., Xu, C., Jiang, T., Initiative, A.D.N., *et al.*, 2010. Abnormal cortical networks in mild cognitive impairment and alzheimer's disease. PLoS computational biology 6, e1001006.

Zhou, M., Zhang, F., Zhao, L., Qian, J., Dong, C., 2016. Entorhinal cortex: a good biomarker of mild cognitive impairment and mild alzheimers disease. Reviews in the Neurosciences 27, 185–195.

Zhu, X., Suk, H.I., Lee, S.W., Shen, D., 2016. Canonical feature selection for joint regression and multi-class identification in alzheimers disease diagnosis. Brain imaging and behavior 10, 818–828.